DOCUMENT RESUME

ED 412 716 EC 305 937

AUTHOR Hoover, John H.; Chalmers, Lynne; Bornfield, Gail
TITLE Assessment of Native American Students under PL 101-476: An Instructional Module for Special Education Courses. Project ERICA.
INSTITUTION North Dakota Univ., Grand Forks. Coll. of Education.; Minot State Univ., ND.
PUB DATE 1997-00-00
NOTE 54p.
PUB TYPE Guides - Classroom - Teacher (052)
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *American Indian Education; *American Indians; *Disability Identification; Elementary Secondary Education; *Emotional Disturbances; Evaluation Methods; Incidence; *Learning Disabilities; Learning Modules; Masters Programs; Student Evaluation; Teacher Education Programs; *Test Bias
IDENTIFIERS *Disproportionate Representation (Spec Educ); Native Americans; North Dakota

ABSTRACT

This module was developed to provide training at the master's degree level for prospective teachers in the areas of learning disabilities and emotional disturbance to serve students in rural and reservation areas of North Dakota. Project ERICA sought to address the overrepresentation of Native American students in North Dakota's special education programs by training graduate students on assessment issues related to the provision of service to Native American students. The module contains an outline of material for infusion into special education assessment courses, a brief narrative section, overhead masters, and a list of suggested readings. Materials address: (1) the overrepresentation of Native American students in special education in North Dakota; (2) the concept of test bias and definitions of test bias; (3) evidence for bias in the assessment of Native American students, including evidence based upon traditional psychometric studies and other evidence; and (4) resolution of test bias by using neuropsychological evaluations, decreasing emphasis on norm-referenced assessments, considering assumptions, using nonverbal tests, using learning style evaluations, developing local norms, examining the opportunity to learn, involving parents and minority team members, and identifying and eliminating biased items. A checklist based on the information reviewed for this module is provided. (CR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 412 716

Assessment of Native American Students Under PL 101-476: An Instructional Module for Special Education Courses

John H. Hoover
and
Lynne Chalmers
The University of North Dakota,
College of Education
and
Human Development
Gail Bornfield
Minot State University
Department of Special Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Module Prepared Under
Federal Grant #CFDA 84.029B, Project ERICA

EC305937

BEST COPY AVAILABLE

ASSESSMENT OF NATIVE AMERICAN STUDENTS UNDER PL 101-476: AN INSTRUCTIONAL MODULE FOR SPECIAL EDUCATION COURSES

Introduction

Project ERICA was funded for three years in 1993 to provide training at the master's degree level for prospective teachers in the areas of learning disabilities and emotional disturbance to serve in rural and reservation areas of North Dakota. A central part of the project was to generate course materials, which would affect favorably upon the delivery of services to children in the area by graduates of the program.

We determined that two aspects of service delivery needed increased attention. At the time the grant was written, the U.S. Office of Civil Rights (North Dakota Advisory Committee, April 27, 1993) had just issued a report that Native American students were significantly overrepresented in North Dakota special education programs. Mission statements issued by both Universities involved in the project (The University of North Dakota and Minot State University) placed service to Native American students as a high priority. We also ascertained that none of the prospective texts for the assessment courses taught under the project's auspices contained sufficient material on assessment issues involving the provision of service to Native American students. Thus, we proposed in the funding request to gather available information and generate course materials related to the assessment of Native American students.

Significant personnel problems exist in rural and remote areas of the United States. The original funding request was based in part on the observation that it is difficult to recruit, train, and retain masters-level professional educators in rural

settings. Further, no widespread agreement exists as to how best to deliver services in such settings, where caseloads may be low and where, resultantly, special education teachers are either itinerant or asked to maintain large, and diverse loads. Thus, a second training module was proposed on the topic of rural service delivery. The modules were developed during the second and third year of the project and were field tested in courses during the third year of funding.

The Native American Assessment module contains the following components: A table of contents, an outline of material for infusion into special education assessment courses, a brief narrative section (constituting lecture notes), overhead masters, and suggested readings.

NATIVE AMERICAN ASSESSMENT AND SPECIAL EDUCATION PLACEMENT

TABLE OF CONTENTS

Outline of Presentations	Page 5
Narrative (Lecture Notes)	Page 6
Nonbiased Assessment Checklist	Page 32
References	Page 35
Overhead Transparency Masters	Page 36, op cit.
Student Versions of Overheads	Page 37, op cit.

PRESENTATION OUTLINE

I. Native American Students in Special Education in North Dakota

- A. Reservations and Tribal Groups
- B. Demographic Considerations
- C. Overrepresentation in Special Education

II. Test Bias

- A. The Concept of Bias
- B. Definitions of Test Bias
  1. Bias based on mean group differences
  2. Regression or line bias
  3. Factor bias
  4. Bias based on consequences of use

III. Evidence for Bias in the Assessment of Native American Students

- A. Evidence based upon traditional psychometric studies
- B. Other evidence and related Issues

IV. Resolving Test Bias

- A. Neuropsychological evaluation
- B. Decreased emphasis on norm-referenced assessment
- C. Consider assumptions
- D. Nonverbal tests
- E. Learning style evaluations
- F. Local norms
- G. Examine opportunity to learn
- H. Involve parents and minority team members
- I. Identify biased items

NARRATIVE (CLASS NOTES)

NATIVE AMERICAN OVERREPRESENTATION IN NORTH DAKOTA

Reservations and Tribal Groups

Four federally-recognized reservations are either wholly or in part located in North Dakota. On an east-west basis, two of the reservations, Turtle Mountain and Fort Totten, are situated centrally, with the Turtle Mountain reservation to the North (see map). Fort Berthold and Standing Rock are further to the west. Standing Rock straddles the border with South Dakota. Information from this section of the document is drawn from Mary Jane Schneider's book, *North Dakota Indians: An Introduction.* The volume is a highly recommended source for educators in North Dakota, virtually all of whom are likely to serve Native American students. This and other useful sources are located in the reference section at the end of this module.

Turtle Mountain. With a population of 9,889 (tribal enrollment = about 26,500; 1990 U.S. Census) Turtle Mountain is the smallest, most densely populated reservation. Most residents can readily get to Belle Court, the town at the geographical and cultural center of the reservation.

Two groups are recognized at Turtle Mountain, the Chippewa (Ojibway) and Michif people. Michif is an Algonquin transliteration of the French word Metis, meaning mixed. Michif people are descended from native people and French fur traders. The language contains both Chippewyan and French words.

Fort Totten. The Fort Totten reservation is located 10 miles south of Devil's Lake, North Dakota. Most of the 2,676 residents of Fort Totten are of Dakota (Sioux)

ancestry. Tribal enrollment was estimated to be 3,900.

Fort Berthold. The Fort Berthold reservation (Pop. 2,999; tribal membership estimated at 9,100) is located in the west-central part of the states straddling Lake Sakakawea. In fact, portions of the reservation were flooded, a matter of great controversy, when the lake was formed in 1954.

Residents of the reservation are members of the three affiliated tribes, the Arikara, Hidatsa, and Mandan. The Garrison diversion project generated considerable hardship on the reservation because farming the rich bottomlands (now under the lake) was central to the cultural and lifestyle of tribal members.

Standing Rock. Standing Rock (Pop. 4,800, of which approximately 2,836 live in North Dakota), straddles the South Dakota Border in western North Dakota; it lies on the west side of Lake Oahe, another Missouri River reservoir which is part of the Garrison Diversion project.

The cultural groups located at Standing Rock are Lakota (Sioux) people, who tend to live on the South Dakota side. Yanktonai is the tribal affiliation of those members living on the North Dakota side of the border.

Demographic Issues

Since the advent of Europeans in North America, the population of Native Americans has declined--through warfare, poverty and systematic bias. Between the 1980 and 1990 census periods, however, a slight increase in North Dakota's native American population has been observed. This is in contrast to a decline (since reversed) in the state's Euro-American numbers (1980 to 1990). Several demographic

issues impact directly on the provision of special education services to Native American students. Primary among these is a tendency toward urbanization.

Native Americans make up approximately 4% of the state's population (25,917/648,800). The fact that Native American population tends to be young, on average, is attested to by the fact that they are estimated to make up nearly 7% of the school-aged population.

Schneider (1994) estimated that the percentage of Native Americans living in North Dakota's four largest cities quadrupled between 1970 and 1990. This trend toward urbanization may be lessened by increased educational, health, and economic opportunities afforded on the reservations, though this remains to be seen. In 1990, about 15% of all North Dakota Native Americans lived in the four largest cities (3,896/25,917). This ran from a low of about 1% of Fargo's population to about 3% in Bismarck (Grand Forks = 2.3%; ,Minot - 2%).

Native American youth are probably drawn to the cities for much the same reason as rural Euro-Americans, e.g., increased education, health care, and economic options are afforded in these locations. The University of North Dakota, for example, draws many Native Americans to Grand Forks. With urbanization, however, some familial supports more readily available in reservation areas are lost.

The urbanization of North Dakota's Native Americans generates significant challenges for North Dakota schools. Cultural and linguistic differences generate specialized educational needs which urban educators may not always be trained to meet. It will be important in the next years to develop close ties and clear communication between reservation schools, tribal officials, and such institutions as

Lake Aggasiz Elementary in Grand Forks (which serves a substantial number of Native American school children). One aspect of services to be considered is whether test bias and assessment procedures in special education tend to unfairly burden Native American children with placement in special education classes.

Overrepresentation Evidence

Overrepresentation exists, according to the federal government, when more Native American (or other minority) students are identified for special education eligibility than would be expected based on their representation in the immediately surrounding population. The excess numbers would have to be more than would be expected by chance--that is to say--statistically significant. In the remainder of this report, we explain how bias in the development, interpretation, and use of tests could produce overrepresentation. In a final section, alternate procedures are outlined.

In April 1993, the North Dakota Advisory Committee to the U.S. Commission on Civil Rights issued a report on the status of Native American students in North Dakota schools special education programs. Serious concerns were expressed regarding the climate of schools, for example, the failure of North Dakota schools to take Native American viewpoints into account during lessons. During the 1991-1992 school year, Native American students accounted for 6.9 percent of students served in public schools. This figure does not include students enrolled in Bureau of Indian Affairs schools.

During the 1991-1992 school year Native American students made up 6.6% of students enrolled in special education, a figure which indicates that statewide Native

American students are not overrepresented in special education. However, in several districts, Native American students appeared on case loads at a higher rate than they appeared in the population. In eleven districts, Native American students were overrepresented, and in twelve districts, fewer Native Americans than expected by their share of the school population were served under PL 101-476. In one district, Bismarck, Native American students were three times as likely to appear on special education course loads than their non-Indian counterparts.

On both a percentage and numerical basis, the most serious overrepresentation was in the Bismarck district, though other districts also showed minority overrepresentation. Representatives of both the state Department of Public Instruction and Bismarck school district have taken measures to reduce this overrepresentation.

TEST BIAS

Bias Defined

Every time assessment practitioners give a test to a child, they are, in effect, taking a momentary snapshot of that youngster's ability, interests, and/or achievement (depending on the purpose of the instrument). Put statistically, this snapshot contains essentially two components; TRUE SCORE (to the extent that the test measures a valid trait reliably) and ERROR (representing randomness).

Random error, slight variations around the TRUE SCORE, is to be expected and is adequately handled by traditional psychometrics via the standard error of measurement and other indices. While a test may be invalid due to a surfeit of random error, this can be addressed by an examination of the reliability coefficient and the

standard error of measurement. Thorndike (1949) provided an excellent list of factors which lead to increased measurement (random error). When error is not random, but is systematic, bias creeps into testing.

One way to conceptualize random error is to suggest that as a practitioner measures a trait via a test or any other observational system, slight variations in scores appear. Perhaps the administrator misunderstands one child and marks a correct answer wrong. It is expected that by way of the randomness of error, the administrator may score the next item too liberally and the errors will balance. Say, however, that the test administrator firmly believes that students short for their age tend to be academically talented. In such a case, perhaps without even realizing it, the administrator systematically scores short students higher and taller students lower. Error is certainly present, but it is not random error. The error illustrated in the scenario is systematic.

Psychometricians refer to such systematic error as test bias and it lies at the heart of all definitions of bias. It is possible that bias against certain groups, or nonrandom systematic error, can result from the test itself, not just the way the practitioner administers, scores, and interprets the instrument. Administrators certainly are a potential source of test bias. Anderson (1981) offered the following useful definition of test bias:

> Bias in a test is a slant in the way a test measures what it is intended to measure; it is a systematic error that disadvantages the test performance of one group compared to another (p. 80).

If a test or other measurement system is designed in such a way that it produces

systematically less meaningful predictions, incorrectly lower scores, or a latent structure different for one identifiable group or another, it is said to be biased. Likewise, if a test user administers, scores, interprets, or puts interpretations to use in a manner to the disadvantage of a linguistic, ethnic, or racial group, the use of the instrument is said to be biased. So, a test, its use and/or interpretation may be biased. In addition, individual items within a test may be biased (Shepard, 1981).

Many test experts distinguish between bias and fairness. Bias refers specifically to the psychometric issue of differences between groups contributing to differential, between group validity. Fairness, on the other hand, refers to the use to which an instrument is put. An unbiased test could potentially be used in a grossly unfair manner, while a slightly biased test could, if interpreted carefully, be put to fair use--at least hypothetically.

Types of Bias

Several lines of evidence are consulted when evaluators and developers consider whether an instrument is biased to the detriment of some group. For reasons which will be explored below, some of these lines of evidence are considered more convincing by psychometricians than others. Among the more commonly recognized types of bias evidence are: (1) mean group differences, (2) regression (or line) bias, and (3) factor bias. A fourth area which has come under consideration recently, closely tied to test validity, may be described as (4) "use" or "outcome of use" bias (Messick, 1988). Each is developed below.

Mean difference bias. Frequently mean differences between groups on an

instrument is cited as evidence for test bias. That is, if members of group B, on average and to a statistically significant degree score lower than members of group A, the test is considered biased against group B.

The logical weakness for considering mean difference as bias is that the question is begged as to whether (1) the differences actually exists, and/or (2) whether the mean difference has, in itself, significant implications for understanding the performance of a single child. Differences are assumed to be nonexistent in nature and thus a manifestation of test bias independently of the instruments use.

For example, a test of arm strength would show mean differences between males and females at age 20, with females scoring lower. Does this mean that the instrument is biased? Not necessarily, unless the test was employed to discriminate by gender on a selection factor, or criterion, unrelated to physical strength. In addition, situations could be imagined where individual differences in strength may be of interest. A test of arm strength which did not "discriminate" between males and females would be suspect on the basis of construct validity.

Nonetheless, in cognitive, motor, achievement, and perceptual tests, it may be most ethical, absent of other data, to assume that observed score differences accrue to tests and are not traceable to characteristics of group identities. At this level, mean differences between groups may, with some trepidation, be taken as evidence for test bias. Mean differences alone would be considered the weakest psychometric evidence for test bias.

At least one federal district court (Larry P. V. Riles) found mean differences in IQ test scores as manifested in unequal placement rates in classes for the educable

mentally handicapped to constitute bias, while another court (PACE v. Hannon) came to the opposite conclusion.

Regression (line) bias. Line bias must be understood in the context of criterion-referenced validity. In criterion-referenced validity (predictive or concurrent being sub-types), the "utility" of a measure is examined via determining whether the test under consideration predicts some useful score. For example, does a college entrance examination predict grades at university?

Regression lines are calculated where test scores are used to predict criteria. Regression solutions are made up of three components, an alpha, which reflects levels of the criterion with "0" amount of the predictor variable; a beta, representing the slope of the prediction line, and an error term. The latter term in regression solutions is taken as a metric of random error and is made up of the average distance of points from a "line of best fit" (Pedhazur, 1986).

Separate regression lines can be calculated for groups of interest to the test developer. If the group-by-regression effect in a statistical test for the regression solution is significant, it means that aspects of the regression solution operated differentially for the two groups. Put another way the instrument is more valid for one group than for another, as it predicts the criterion better for one group than for the other. Of course, as described here, it is assumed that the criterion is one which practitioners would wish to be predicted via the test in question!

Essentially three possibilities for line bias exist. First, the two groups could achieve different alpha levels with parallel regression lines (that is, equal betas). In this case, a single line averaged across both groups would over-predict the

performance of one group on the criterion (based on test performance) and under-predict the other. The second possibility is that the regression lines are not parallel; put another way, the betas are not equal. In such an instance, an instrument would predict the criterion "less accurately" for one group than for the other. Under extreme conditions, it is even possible that the test is valid for group A, but shows a flat regression line (that is, does not predict the criterion) for group B. Finally, in a third possibility, the test might show both unequal betas and alphas for the groups under consideration.

Factor Bias. Factor or component analysis (hereafter FA) refers to a collection of mathematical methods for examining relationships between items and composites on an instrument or across two or more tests. Often, FA is employed by test developers to determine the pattern of an instrument's scores and/or to compare the structure of a test with some theoretical view of a trait under consideration.

For example, a theory might have been promulgated suggesting that personality is divisible into two relatively independent dimensions, X and Y. Independent in this case means that a person's score on dimension X and his/her score on dimension Y are not necessarily related. Using FA, an investigator may be able to determine whether an instrument designed around the X-Y theory "behaves" as the X-Y model predicts it will; that is, that items "factor" into two relatively independent structures, corresponding to X and Y as predicted by the theorist. Of course, FA can also be used to examine a test's structure independently of a theoretical orientation.

Factor analysis can be performed on tests separately between two or more

groups. If the structure of the instrument is different between the groups, then interpretation is exceedingly problematic. Thus, it could be said that factor bias exists when the instrument is used on the group where it does not factor as predicted or desired.

Using the example cited above, it is possible that for group A the ACME Test or Personality behaves as it should, given the X-Y theory on which it is based. That is items factor into a structure where they are related either to factor (personality trait) X or factor Y (a second personality trait). These two factors are relatively independent of one another. Perhaps for a second group, B, however, items factor into three factors, G, H, and I. It is even possible that none of the three factors seen by Group B subjects is similar to factors X and Y, as scored by group A.

Use bias. In considering validity of use of an instrument, particularly within a specific cultural context, Messic (1988) suggested that users consider a dimension which he called, "implications of use." This means that the implications of using an instrument and of embracing the theory on which it is based, should be considered as part of the validity package.

Even an ostensibly valid and reliable test of general cognitive ability is based upon a particular theory of human ability. Thus, use of the instrument implies acceptance of the theory and an understanding of the possible results of accepting this view. In the case of general ability measures, one outcome of use is the possibility that the student being assessed will be declared to be mentally retarded under PL 101-476. If a practitioner views this an undesirable, he or she may reconsider use of the instrument in light of this concern. Indeed, several researchers have argued that the

"dividing out" of students at all on the basis of norm referenced testing is, in and of itself, anti-egalitarian and therefore biased. On the other hand, other writers have pointed out that unlabeled students do not receive services, a problem perhaps more severe in some cases than those associated with normative testing (MacMillan, 1991).

EVIDENCE REGARDING TEST BIAS AGAINST NATIVE AMERICAN STUDENTS

Studies Based on Traditional Psychometrics

In an ERIC and Psychlit search, many articles and papers were located which dealt with testing of Native American youngsters and adults. Commonly, these papers reflect the view that assessment bias is one reason that Native American students are overrepresented in special education (e.g., Neely & Shaughnessy, 1989). Unfortunately, from a technical perspective, not many data-based studies turned up, e.g., few studies were found which dealt with factor or line bias.

Writers citing differential levels of placement, are invoking mean bias. Again, the problem with this definition of bias is that legitimate concerns in performance which derives from cultural or socioeconomic differences may be ignored if mean score differences are considered bias. Another way to think of this is to pose the following questions: Is a differential rate of special education placement, in and of itself, evidence for bias in general or test bias in particular? Is it possible that differential placement rates between Native American and Anglo students represent a natural response of the system resulting from differing socioeconomic pressures on rates of learning?

In personal communication with a school psychologist in Belle Court, it was determined that students on the Turtle Mountain reservation demonstrated about a 1/3 or 5 pts. of a SD discrepancy from Euro-American scores. As stated above, this constitutes weak, though noteworthy, evidence for bias. According to these data, nearly all of the norm-referenced achievement and cognitive ability instruments

typically employed in special education eligibility determination and programming decisions should be interpreted cautiously for this population.

We were able to locate one study of line bias on a kindergarten screening battery (Stone & Gridley, 1991). Tests evaluated across groups were the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981; The Kindergarten Language Screening Test (Madison & Gauthier, 1982), the McCarthy Draw-A-Child Test (McCarthy, 1972), and the Beery Developmental Test of Visual-Motor Integration (Beery, 1982). The criterion selected was for the Stanford Achievement Test (SAT) - Total Battery Score (Psychological Corporation, 1983). The predictor variables were administered during students' kindergarten year and used to predict the SAT Total Battery at first and second grade. The predictor battery was correlated with performance of both groups at both first and second grade. Performance of Native American students were slightly overestimated by the kindergarten battery when compared to the prediction equation for European-American students. Substantial mean differences were noted for both the tests in the kindergarten battery and the criterion test. This does not constitute evidence for bias against the Native American students.

Mishra (1982) examined item bias in portions of the WISC-R among Navajo children. Fifteen of 73 items on the information, similarities, and vocabulary sub-tests were biased against the Navajo students. This means that the items were differentially more difficult within ability levels. Of the three tests, vocabulary produced the highest percentage of biased items. Some evidence for caution in the interpretation of ability measures is warranted based on this study. Similar results were reported by Ross-

Reynolds and Reschley (1983), though the effects were generally weaker. No studies were found dealing with any of the North Dakota groups specifically.

Davidson (1989) found that gifted and talented Native American students scored differently than their non-Indian counterparts on the K-ABC. Anglo students scored highest on sequential processing, while the Native American students scored higher on simultaneous processing. No evidence accrued that the test factored differently for the two groups, however.

At least as traditionally conceptualized, little evidence exists for widespread test bias in the types of instruments used in special education placement and eligibility decisions. Information-heavy subtests on such instruments as the WISC-III show a small, but significant amount of item bias, suggesting that ability tests should be interpreted cautiously. Evidence from Davidson's study of gifted young people suggests that practitioners look into use of the K-ABC simultaneous processing score as the best measure of ability among Native Americans. Unfortunately, the simultaneous processing score on the K-ABC shows relatively low predictive validity. Very few studies examined the types of achievement tests typically administered by special educators.

Conclusions Regarding test Bias

An examination of the research literature suggests the following conclusions:

1. Ability and achievement tests may show a 1/2 SD difference between Native American students in North Dakota attending schools on reservations and the norming sample. Tentatively, until other information is made available, this

could be considered evidence of test bias.

2. Little evidence has accrued suggesting that the typical ability and achievement tests used in special education assessment predict criteria in a manner disadvantageous to Native American Students.
3. Not much evidence is available suggesting that the typically-employed instruments are structured significantly differently from each other. In one study it was shown that Native Americans show significantly higher functioning in simultaneous processing when compared with Euro-Americans on the K-ABC.
4. Differential item-difficulties between Native American groups and Euro-Americans have been shown. Test users should examine items carefully for content unfamiliar to their students.

Other Sources of Inappropriate Placement

Learning Styles and Cultural Explanations for Bias

Sternberg (1995) argued that assessment practices in the United States concentrate on only a portion of the abilities important to learning. Aside from the analytical "type" of intelligence stressed in Western cultures, at least two more abilities are hypothesized: creative and practical intelligence. Sternberg and his colleagues have convincingly demonstrated that instruction geared to these nontraditional types of abilities succeed with students high in measures of these intelligences. Other writers have suggested a variety of models for human ability beyond the global intellectual model which has dominated psychology for the past 50 years. For example, Guilford posits over 120 "intelligences." It is beyond the scope of this module to explore these

theories in depth, but it is important to consider that Native American children may be different enough in learning style that they both score low on tests and achieve lower in school because schools favor the analytical learning style represented by traditional measures of intelligence.

Experts acknowledge individual variability and differences among the over 300 tribes in North America (and approximately 200 languages) (Locke, et al., 1977). Nonetheless, some important differences between the learning styles of Native American and Anglo children have been posited which may affect selection, administration, and interpretation of test instruments to children from the tribes.

Non-Competitiveness

Many Native American cultures eschew individual accomplishments in favor of group achievement (Neely & Shaughnessy, 1989). If a Native American child being assessed comes from such a cultural background, they may not perform up to their own capabilities on an individual test. Such less than optimal performance would, of course, make interpretation difficult.

Language

We were unable to locate data indicating the percentage of students who speak one of the Native American languages. In fact, calls to reservation schools indicate that nearly all students currently receiving services use English as their first and primary language. However, many students, both Native and Euro-Americans who live in straightened circumstances in rural areas come to school (and to testing situations)

with restricted expressive language. It is important that test administrators not mistake language differences, dialects, nor even disabilities with other disabilities such as mental retardation.

Most Native American students have heard grandparents speak their language of origin. In addition, to some extent in virtually all reservation schools, native languages are being offered as part of cultural and self-esteem programming. Test administrators may hear a smattering of Ojibway, French, Arikara, or other languages when they assess students.

Thinking Styles

We have long recognized the intimate relationship between language and thought. In addition, many fundamental differences in cognition have been noted across cultures. It is possible that Native American children, even those whose primary language is English, process information in ways which are both different and which do not coincide well with standardized testing.

Kaufman and Kaufman (1983) argued that the type of logic-deductive, sequential reasoning favored in schools and by tests is a phenomenon of Western culture which may be alien to many children in our increasingly-diverse schools. In fact they have developed a test around this notion called The Kaufman Assessment Battery for Children (Kaufman & Kaufman, 1983; K-ABC) which may be a more reasonable test of cognitive performance for Native American students, particularly the "simultaneous processing" score. It is important to note that limited research has accrued on this possibility. The K-ABC has been proposed as a better method for

tapping the intellectual abilities as opposed to disabilities of Native American students (Davidson, 1989). Much more research is needed in this area.

The K-ABC has been shown to be the intellectual ability measure with the least predictive validity (Ysseldyke & Algozzine, 1993). However, this may be because American schools and traditional IQ measures are on the same analytical wavelength (Sternberg, 1994). Perhaps as teachers become more adept at adjusting to various learning styles, they will be able to respond well to Kaufman's simultaneous processors.

Several other possible learning style issues were developed in an excellent ERIC monograph (Swisher, 1991). Understanding that listing such characteristics runs the risk of stereotyping, they are addressed below:

Learning by observation. Several studies have shown that members of plains tribes tend to prefer an observation-in-practice format for learning. This has also been observed among Southwestern groups such as the Navajo and Yaqui.

Hesitant to speak in large or small groups. Some Native American children prefer one-to-one learning conversations with teachers to speaking in class-wide or small groups. The latter characteristic differs by cultural group and may be more common among Dakota Sioux children than among Ojibway or especially Michif youngsters.

Examinee or Use Issues

Several writers have addressed problems when majority practitioners administer tests to minority students. These were summarized by Reynolds (1982). We

have paraphrased them below in light of the assessment of Native American students:

Inappropriate content. The items and content of the test are grossly unfamiliar to Native American students. For example, the content of the test clearly is geared toward middle class values--over and above that required by the content of the test. For example, many Indian children live in rural areas while test content is often geared to urban environments.

Inappropriate standardization samples. Seldom are Native American students represented in norming samples. In such cases, practitioners must interpret cautiously. Larger districts with many Native American students or districts serving primarily students from tribal backgrounds may wish to consider developing local norms for tests which they deem otherwise useful

Examiner and language bias. Ability and achievement levels may be systematically underestimated if even subtle linguistic differences exist between administrators and examinees. This makes the development of rapport an essential component of the evaluation and identification process. In addition, cultural differences between the expectations of students and examiners may produce inaccurate results (Neely & Shaaughnessy, 1989). Native American students may feel intimidated by the testing situation itself, though this factor is better established for African American students.

Inequitable social consequences. The above-mentioned biasing effects may produce unfair social outcomes. This factor is clearly a reference to the unfair use of tests. It may be, for example, contrary to a districts policy to allow overrepresentation to occur beyond a certain point.

Solutions and Practices

Neuropsychological Evaluations

In that special education evaluation is designed in part to determine whether students are truly disabled and not culturally different, it has been proposed that more attention be paid to neurological evaluations. Progress in this area has been accelerating and practitioners should make themselves aware of advances in the neurological domain. As they evolve, neuropsychological evaluations should be less prone to misidentify children as being disabled whose behavioral differences are cultural in nature. A danger in relying only on biological indicants of disability is that some students who need services will not receive them.

Decreased Emphasis on Psychometrics

Leung (1996) noted that both informal direct methods and clinical judgement may be superior to traditional psychometric assessment where some concern exists that a student may be labeled as disabled because he or she belongs to a minority group. The University of Minnesota Learning Disabilities Institute has proposed methods whereby divergence from curriculum-based assessment measures may be employed as opposed to psychometrics for determining eligibility. Because students are identified as disabled based on divergence from data collected in their home schools, this amounts to informal adoption of local norms. These methods have proved reliable as well as valid.

Interviews, checklists, adaptive behavior scales, and direct observation are all

methods beyond testing that may yield data which may be used by interdisciplinary team members to make informed decisions about eligibility and programming for Native American students.

Learning Style Evaluations

Several methods for evaluating students learning styles rather than only their divergence from norms have been proposed. No mechanism has been proposed for using such information in lieu of traditional eligibility data, so at present, assessments based on learning style would be a supplement to special education placement not an alternative (Swisher, 1991). Data for learning style types of evaluations is only, minimally available, at this time.

Local Norms

Districts serving many Native Americans could establish norms for such individuals and use these to determine eligibility. To the best of our knowledge, about a 1/2 standard deviation difference exists in North Dakota, though this is not well established. Thus, eligibility for mental retardation services would run from IQ 65 and down, for example. Expected scores on the WISC-R would be 95 rather than 100. No work has been done with WISC-III. Presumably, fewer students would score in the discrepancy ranges required for LD eligibility.

Examine Opportunity to Learn

Students should not be labeled as disabled when their learning difficulties can

better be explained by lack of an opportunity to learn. Thus, students from deprived environments who missed a great deal of school would receive interventions based on these needs rather than special education services. Eligibility requirements for the PL 101-476 LD and SED categories both include exclusions based on educational deprivations.

Consider Assessor's Assumptions

Each evaluator who sees Native American children in their practice should pause to consider whether their own biases may enter into administration, scoring, interpretation, or decision making. It has been shown that expectations for success or failure can alter test scores by small but significant amounts. Given tester effects, it is also important to work much harder on rapport with a student who may be ill at ease due to racial and/or cultural differences.

Involve Parents/Community

Leung (1996) pointed out that fewer biased decisions are made when assessors consider the opinions of parents and members of the community representing the child being evaluated. Parents would be likely to relate, if interviewed sensitively, whether the youngster diverges not just from Euro-American norms, but from the age and culture-based expectations of the family and community. A students who diverges from these norms is likely to be disabled and not just culturally different.

Nonverbal Tests

When questions arise about whether a student's observed general cognitive ability result may be due to language problems, the psychologist may wish to consider employing nonverbal measures. Several of these, including the Arthur Adaptation of the Leiter, the Test of Nonverbal Intelligence, and the Psychoeducational Profile are available.

Involve Minority Team Members

In situations where bias may be producing overrepresentation or where this may be a concern, it may be appropriate to involve minority professionals as advocates for students.

Identifying Bias in Test Items

Test publishers have promulgate standards, based on expert analysis, to prevent items that may be offensive to members of minority groups from appearing on tests. According to Shepard (1981), elements of this analysis include the avoidance of biased language, proportionate representation of minorities in test items, and avoidance of offensive stereotypes in the manner in which minorities are portrayed.

The representation of Native American people and culture in items may be problematic because depiction rates are based on population proportion and Native Americans represent a relatively small proportion of the population. We are not familiar with any tests which specifically portray reservation or Native American city life in items or examples. However, due to the above-mentioned standards, children and adults portrayed in such instruments as the Woodcock Psychoeducational Battery-Revised

(Woodcock & Johnson, 1989) and the Wechsler Individual Achievement Test do appear more diverse than in older instruments. One drawback to judgement of item bias is that experts have been unable to successfully predict difficulty based on supposedly biased item content (Sandoval & Miller, 1980).

The sensitive practitioner should not assume that items are not biased, despite the standards adopted by publishers. When examining test materials, psychologists and special educators ought to examine them for stereotyped portrayals, linguistic patterns which conflict with local norms, and a lack of portrayal of persons of color in items. Hambelton (1980) proposed a list of questions (paraphrased in Shepard, 1981) which may prove useful to practitioners working with Native American children:

1. Is the item free of offensive sexual, cultural, racial, and/or ethnic content?
2. Is the item free of sexual, cultural, racial, and/or ethnic stereotyping?
3. Is the item free from language that could be offensive to a segment of the examinee population?
4. Is the item free from descriptions that could be offensive to a segment of the examinee population?
5. Will activities or situations described in the item be familiar to all examinees?
6. Will the words in the item have a common meaning to all examinees?
7. Is the item free from difficult vocabulary and/or sentence structure?
8. Will the item format be familiar to all examinees?

We would add the following question related to the instrument and individual items: Overall, are persons of color adequately represented in test materials?

Statistical methods for examining item bias are also available. Most are based

on an examination of item difficulty statistics between groups. If too many items are thus biased against Native American students, an alternative test could be selected. Note that practitioners should not attempt to modify or delete items based on judgement or statistical item bias indices as this practice will invalidate results.

Consider scoring alternates in the following circumstances:

1. When the response is correct given local semantics--word meanings, it should be scored correctly. For example "tuke" for "stocking cap".
2. In language tests, score syntax as correct if it matches local standards for word order. This will require a considerable amount of expertise on the part of the evaluator.

A checklist based on the information reviewed for this module is provided. This may be reproduced by users of the module.

A CHECKLIST FOR NONBIASED ASSESSMENT

1. Language differences have been considered.

The child's primary language is: _____

____ No evidence of impact on English from other language(s).

____ Assessment for a language disorder has been undertaken. Results:

_____

_____

_____

_____

Note: If learning problem results from a language disorder, avoid using eligibility under LD, SED, or EMH unless compelling reasons exist to suspect one of these disabilities. List reason here:

_____

_____

_____

2. Cultural differences have been considered.

____ Tester effects not present (based on child's reaction to evaluators).

____ Non-Native American evaluators have considered preconceived notions about youngsters being assessed.

____ Parents/guardians/interested extended family members have been interviewed and agree that the child's cognitive, emotional, academic, or behavioral status differs from culturally-expected norms. Specify:

_____

_____

_____

_____

____ Native American educational professionals were included on decision-making team or data were reviewed by a Native American expert on special education decision-making.

____ No evidence exists that the suspected disability results from the interaction of Native American and Euro-American culture.

3. The opportunity to learn has been considered.

____ The student has not missed an excessive number of school days in the past one year (not to exceed 20 days).

____ The student has not transferred between schools more than one time during the past one year.

____ The educators primarily involved with the student are aware of possible individual and culture-based learning style differences and have adapted instruction accordingly. Evidence:

_____

_____

_____

4. Local (or culturally-specific) norms were employed in interpreting cognitive ability and/or academic achievement tests.

Note: It is recommended that, unless a majority of blanks are checked and the evidence is convincing, students be served under title, ESL, or general

education categories, not special education. If any evidence exists of expressive or receptive language differences or deficits, serve under speech/language, avoiding use of the EMH, LD, or SED labels. Please list below why LD, EMH, or SED label has been applied to the child:

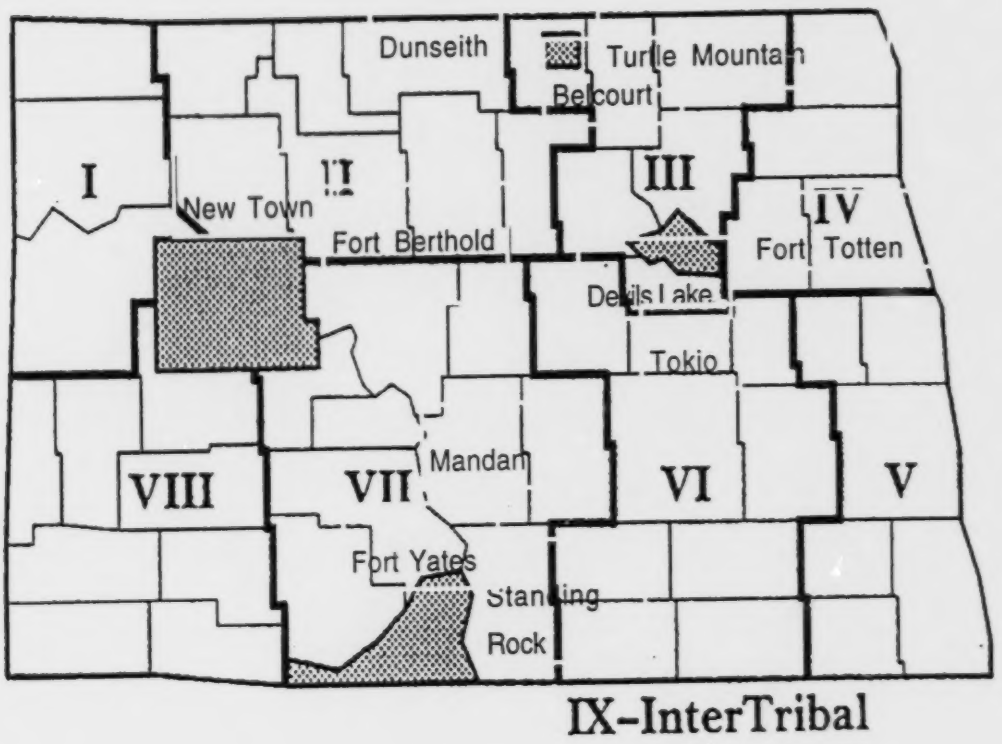
_____

_____

_____

_____

Figure Example 1. Map showing North Dakota Indian reservations and related communities.



BEST COPY AVAILABLE

REFERENCES

Carney, L.J., & Chermak, G.D. (1991). Performance of American Indian children with fetal alcohol syndrome on the test of language development. Journal of Communication Disorders, 24, 123-124.

Davidson, K.L. (1992). A comparison of Native American and white students' cognitive strengths as measured by the Kauffman Assessment Battery for Children. Roeper Review, 14, 111-115.

Leung, B.P. (1996). Quality assessment practices in a diverse society. Teaching Exceptional Children, 23, 42-45.

Neely, R. & Shaughnessy, M.F. (undated). Assessment and the Native American. (ERIC document reproduction service, No. ED 273 889).

Schneider, M.J. (1994) North Dakota Indians: An Introduction. Dubuque, IA: Kendall/Hunt.

Stone, B. & Gridley, B. (1991). Test bias of a kindergarten screening battery: Predicting achievement for white and Native American elementary students, School Psychology Review, 20, 132-139.

Swisher, K. (1991). American Indian/Alaskan Native learning styles: Research and Practice. Washington, D.C.: Report of the Office of Educational Research and Improvement. (ERIC document reproduction service, No. ED 333 175).

Tests appropriate for use with American Indians. (1977). Princeton, NJ: Educational testing service test collection. (ERIC document reproduction service, No. ED 213 546).

Tonemah, S. (1987). Assessing American Indian gifted and talented students' abilities. Journal for the Education of the Gifted, 10, 181-194.

OVERHEAD TRANSPARENCY MASTERS

Project ERICA

Demographic Trends

- The Native American Population was 4% of the ND total in 1990 (25,917/648,800).

- The percentage of the population living in the four largest cities quadrupled between 1970 and 1990.

- The population is relatively young: 4% of overall population, nearly 8% of school-age population.

Project ERICA

Reservations and Tribal Groups

△ TURTLE MOUNTAIN

- Chippewa
- Michif (Metis)

△ FORT TOTTEN

- Dakota (Sioux)
- Yanktonai

△ FORT BERTHOLD

- Arikara
- Hidatsa
- Mandan

△ STANDING ROCK

- Lakota
- Yanktonai

Project ERICA

Urbanization

CENTER	NAT AMER#	TOTAL N	%NAT AMER
Bismarck	1,261	49,256	2.6
Fargo	796	74,111	1.1
Grand Forks	1,115	49,425	2.3
Minot	724	35,544	2.0

3,896 of 25,917 Native American people (15%) lived in North Dakota's four largest cities as of 1990.

Project ERICA

Districts in which American Indian students are over-represented

- Fort Totten (76% served v. 100% in SPED)
- Bismarck (3.8% v. 8.4%)
- School for the Blind (22.2% v. 33.3)
- Dickey LaMoure (0.5% v. 2.3%)
- Morton (17.8% v. 27.3%)
- Lake Region (8.8% v. 13.2%)
- Ann Carlson (7.8% v. 10.0%)
- Griggs/Steel/Trail (0.5% v. 1.7%)
- Upper Valley (0.5% v. 1.6%)
- State Developmental Center (16.7% V. 25%)
- Souris Valley (6.6% V. 9.7%)

Project ERICA

Test Bias

DEFINITION:

Bias . . . is a slant in the way a test measures what it is intended to measure; it is a systematic error that disadvantages the . . . performance of one group as compared to another (Shepard, 1981).

TYPES OF BIAS:

1. Mean Group Differences
2. Regression or Line Bias
3. Factor Bias
4. Implications of Use Bias

Project ERICA

Bias Based on Mean Group Differences

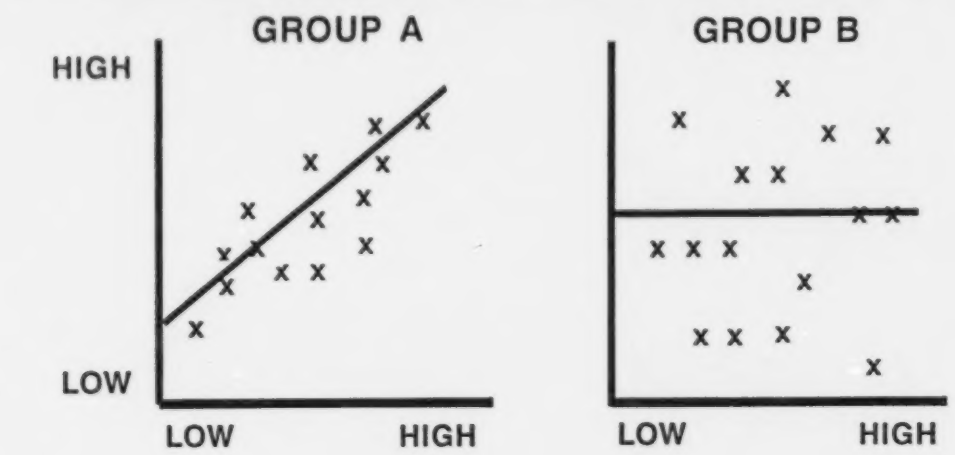
A test may be declared to be biased if groups A and B achieve statistically significant mean score differences on the instrument.

- Weakest form of bias evidence
- Assumes observed difference is attributable to test
- Safest assumption, absent other data, is to assume that racial differences on test performance constitutes bias

Project ERICA

Regression or Line Bias

LINE BIAS refers to a test's differential prediction of a criterion between two racial, ethnic, or linguistic groups.



TEST SCORES

For group A, the test accurately predicts the criterion.

For group B, it does not. The test is biased against group B.

The line shown in the figures is a regression or prediction line.

Project ERICA

Factor Bias

FACTOR BIAS. . .

refers to differences in test structure between two groups.

The test "behaves" differently for members of the two groups.

Project ERICA

Use Bias

USE BIAS . . .

Means the outcome of use of the test or the theory on which it is based is disadvantageous to one or more groups

Project ERICA

Cultural Differences which may affect Test Administration & Interpretation

- Non Competitiveness
- Language
- Cognitive Style
- Learning by Observation
- Hesitant to speak in groups

Project ERICA

Other Problems with Norm-Referenced Assessment

- Inappropriate Content
- Inappropriate Standardization Sample
- Examiner and Language Effects

Project ERICA

Solutions and Practices

- Consider Neuropsychological Evaluation
- Decrease Emphasis on Traditional Psychometrics
- Consider Assumptions
- Nonverbal Tests
- Learning Style Evaluation
- Examine Opportunity to Learn
- Involve Minority Team Members and Parents
- Identification of Bias in Items

STUDENT VERSIONS OF OVERHEADS

Project ERICA — Demographic Trends

- The Native American Population was 4% of the ND total in 1990 (25,917/648,800).
- The percentage of the population living in the four largest cities quadrupled between 1970 and 1990.
- The population is relatively young: 3% of overall population, nearly 8% of school-age population.

Project ERICA — Reservations and Tribal Groups

- Δ TURTLE MOUNTAIN
  - Chippewa
  - Michif (Metis)
- Δ FORT TOTTEN
  - Dakota (Sioux)
  - Yanktonai
- Δ FORT BERTHOLD
  - Arikara
  - Hidatsa
  - Mandan
- Δ STANDING ROCK
  - Lakota
  - Yanktonai

Project ERICA — Urbanization

CENTER	NAT AMER#	TOTAL N	%NAT AMER
Bismarck	1,261	49,256	2.6
Fargo	796	74,111	1.1
Grand Forks	1,115	49,425	2.3
Minot	724	35,544	2.0

3,896 of 25,917 Native American people (15%) lived in North Dakota's four largest cities as of 1990.

Project ERICA — Districts in which American Indian students are over-represented

- Fort Totten (76% served v. 100% in SPED)
- Bismarck (3.8% v. 8.4%)
- School for the Blind (22.2% v. 33.3)
- Dickey LaMoure (0.5% v. 2.3%)
- Morton (17.8% v. 27.3%)
- Lake Region (8.8% v. 13.2%)
- Ann Carlson (7.8% v. 10.0%)
- Griggs/Steel/Trail (0.5% v. 1.7%)
- Upper Valley (0.5% v. 1.6%)
- State Developmental Center (16.7% V. 25%)
- Souris Valley (6.6% V. 9.7%)

BEST COPY AVAILABLE

Project ERICA — Test Bias

DEFINITION:

Bias . . . is a slant in the way a test measures what it is intended to measure; it is a systematic error that disadvantages the . . . performance of one group as compared to another (Shepard, 1981).

TYPES OF BIAS:

1. Mean Group Differences
2. Regression or Line Bias
3. Factor Bias
4. Implications of Use Bias

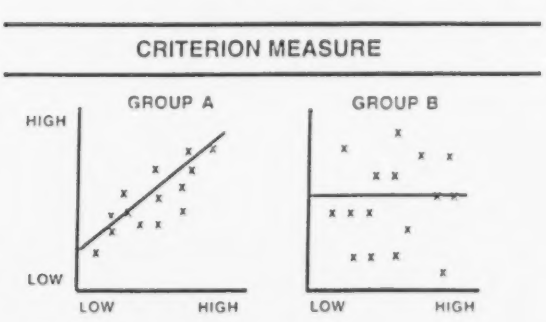
Project ERICA — Bias Based on Mean Group Differences

A test may be declared to be biased if groups A and B achieve statistically significant mean score differences on the instrument.

- Weakest form of bias evidence
- Assumes observed difference is attributable to test
- Safest assumption, absent other data, is to assume that racial differences on test performance constitutes bias

Project ERICA — Regression or Line Bias

LINE BIAS refers to a test's differential prediction of a criterion between two racial, ethnic, or linguistic groups.



For group A, the test accurately predicts the criterion. For group B, it does not. The test is biased against group B. The line shown in the figures is a regression or prediction line.

Project ERICA — Factor Bias

FACTOR BIAS . . .

refers to differences in test structure between two groups.

The test "behaves" differently for members of the two groups.

BEST COPY AVAILABLE

Project ERICA — Use Bias

USE BIAS . . .

Means the outcome of use of the test or the theory on which it is based is disadvantageous to one or more groups

Project ERICA — Cultural Differences which may affect Test Administration & Interpretatio

- Non Competitiveness
- Language
- Cognitive Style
- Learning by Observation
- Hesitant to speak in groups

Project ERICA — Other Problems with Norm-Referenced Assessment

- Inappropriate Content
- Inappropriate Standardization Sample
- Examiner and Language Effects

Project ERICA — Solutions and Practices

- Consider Neuropsychological Evaluation
- Decrease Emphasis on Traditional Psychometrics
- Consider Assumptions
- Nonverbal Tests
- Learning Style Evaluation
- Examine Opportunity to Learn
- Involve Minority Team Members and Parents
- Identification of Bias in Items